

WEST Search History

Hide Items

Restore

Clear

Cancel

DATE: Thursday, May 20, 2004

| Hide?                    | Set Name  | Query                    | Hit Count |
|--------------------------|---|--------------------------|-----------|
|                          | <i>DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI; PLUR=YES; OP=ADJ</i> |                          |           |
| <input type="checkbox"/> | L16   | MEGF10                   | 1         |
| <input type="checkbox"/> | L15   | human EGF                | 909       |
| <input type="checkbox"/> | L14   | L13 AND MEGF10           | 0         |
| <input type="checkbox"/> | L13   | L12 AND human            | 2057      |
| <input type="checkbox"/> | L12   | L11 AND EGF              | 2067      |
| <input type="checkbox"/> | L11   | 536/23.1,23.4,23.5.CCLS. | 17622     |
| <input type="checkbox"/> | L10   | L9 AND MEGF10            | 1         |
| <input type="checkbox"/> | L9  | L8 AND human             | 2819      |
| <input type="checkbox"/> | L8  | L7 AND EGF               | 2828      |
| <input type="checkbox"/> | L7  | 435/325.CCLS.            | 14916     |
| <input type="checkbox"/> | L6  | Maranda-M.IN.            | 1         |
| <input type="checkbox"/> | L5  | Miranda-Maricar.IN.      | 22        |
| <input type="checkbox"/> | L4  | Miranda.IN.              | 911       |
| <input type="checkbox"/> | L3  | Yu-X.IN.                 | 546       |
| <input type="checkbox"/> | L2  | Yu-Xuanchuan.IN.         | 22        |
| <input type="checkbox"/> | L1  | (Yu.IN.)                 | 87877     |

END OF SEARCH HISTORY

Connecting via Winsock to STN  
Welcome to STN International! Enter x:x  
\*\*\*\*\* Welcome to STN International \*\*\*\*\*  
\*\*\*\*\* STN Columbus \*\*\*\*\*

FILE 'HOME' ENTERED AT 11:30:52 ON 20 MAY 2004

=> file BIOSCIENCE

FILE 'DRUGMONOG' ACCESS NOT AUTHORIZED

FILE 'ADISCTI' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 Adis Data Information BV

FILE 'ADISINSIGHT' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 Adis Data Information BV

FILE 'ADISNEWS' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 Adis Data Information BV

FILE 'AGRICOLA' ENTERED AT 11:31:02 ON 20 MAY 2004

FILE 'ANABSTR' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 THE ROYAL SOCIETY OF CHEMISTRY (RSC)

FILE 'AQUASCI' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT 2004 FAO (On behalf of the ASFA Advisory Board). All rights reserved.

FILE 'BIOBUSINESS' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 Biological Abstracts, Inc. (BIOSIS)

FILE 'BIOCOMMERCE' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 BioCommerce Data Ltd. Richmond Surrey, United Kingdom. All rights reserved

FILE 'BIOSIS' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 BIOLOGICAL ABSTRACTS INC.(R)

FILE 'BIOTECHABS' ACCESS NOT AUTHORIZED

FILE 'BIOTECHDS' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 THOMSON DERWENT AND INSTITUTE FOR SCIENTIFIC INFORMATION

FILE 'BIOTECHNO' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 Elsevier Science B.V., Amsterdam. All rights reserved.

FILE 'CABA' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 CAB INTERNATIONAL (CABI)

FILE 'CANCERLIT' ENTERED AT 11:31:02 ON 20 MAY 2004

FILE 'CAPLUS' ENTERED AT 11:31:02 ON 20 MAY 2004

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2004 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'CEABA-VTB' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (c) 2004 DECHEMA ev

FILE 'CEN' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 American Chemical Society (ACS)

FILE 'CIN' ENTERED AT 11:31:02 ON 20 MAY 2004

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2004 American Chemical Society (ACS)

FILE 'CONFSCI' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 Cambridge Scientific Abstracts (CSA)

FILE 'CROPB' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'CROPU' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'DISSABS' ENTERED AT 11:31:02 ON 20 MAY 2004

COPYRIGHT (C) 2004 ProQuest Information and Learning Company; All Rights Reserved.

FILE 'DDFB' ACCESS NOT AUTHORIZED

FILE 'DDFU' ACCESS NOT AUTHORIZED

FILE 'DGENE' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'DRUGB' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'DRUGMONOG2' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 IMSWORLD Publications Ltd

FILE 'IMSDRUGNEWS' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 IMSWORLD Publications Ltd

FILE 'DRUGU' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'IMSRESEARCH' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 IMSWORLD Publications Ltd

FILE 'EMBAL' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Elsevier Inc. All rights reserved.

FILE 'EMBASE' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Elsevier Inc. All rights reserved.

FILE 'ESBIOBASE' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Elsevier Science B.V., Amsterdam. All rights reserved.

FILE 'FEDRIP' ENTERED AT 11:31:02 ON 20 MAY 2004

FILE 'FOMAD' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Leatherhead Food Research Association

FILE 'FOREGE' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Leatherhead Food Research Association

FILE 'FROSTI' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Leatherhead Food Research Association

FILE 'FSTA' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 International Food Information Service

FILE 'GENBANK' ENTERED AT 11:31:02 ON 20 MAY 2004

FILE 'HEALSAFE' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Cambridge Scientific Abstracts (CSA)

FILE 'IFIPAT' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 IFI CLAIMS(R) Patent Services (IFI)

FILE 'IMSPRODUCT' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 IMSWORLD Publications Ltd

FILE 'JICST-EPLUS' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Japan Science and Technology Agency (JST)

FILE 'KOSMET' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 International Federation of the Societies of Cosmetics Chemists

FILE 'LIFESCI' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Cambridge Scientific Abstracts (CSA)

FILE 'MEDICONF' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (c) 2004 FAIRBASE Datenbank GmbH, Hannover, Germany

FILE 'MEDLINE' ENTERED AT 11:31:02 ON 20 MAY 2004

FILE 'NIOSTIC' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 U.S. Secretary of Commerce on Behalf of the U.S. Government

FILE 'NTIS' ENTERED AT 11:31:02 ON 20 MAY 2004  
Compiled and distributed by the NTIS, U.S. Department of Commerce.  
It contains copyrighted material.

All rights reserved. (2004)

FILE 'NUTRACEUT' ENTERED AT 11:31:02 ON 20 MAY 2004  
Copyright 2004 (c) MARKETLETTER Publications Ltd. All rights reserved.

FILE 'OCEAN' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Cambridge Scientific Abstracts (CSA)

FILE 'PASCAL' ENTERED AT 11:31:02 ON 20 MAY 2004  
Any reproduction or dissemination in part or in full,  
by means of any process and on any support whatsoever  
is prohibited without the prior written agreement of INIST-CNRS.  
COPYRIGHT (C) 2004 INIST-CNRS. All rights reserved.

FILE 'PCTGEN' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 WIPO

FILE 'PHAR' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 PJB Publications Ltd. (PJB)

FILE 'PHARMAML' ENTERED AT 11:31:02 ON 20 MAY 2004  
Copyright 2004 (c) MARKETLETTER Publications Ltd. All rights reserved.

FILE 'PHIC' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 PJB Publications Ltd. (PJB)

FILE 'PHIN' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 PJB Publications Ltd. (PJB)

FILE 'PROMT' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Gale Group. All rights reserved.

FILE 'RDISCLOSURE' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Kenneth Mason Publications Ltd.

FILE 'SCISEARCH' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT 2004 THOMSON ISI

FILE 'SYNTHLINE' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 Prous Science

FILE 'TOXCENTER' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 ACS

FILE 'USPATFULL' ENTERED AT 11:31:02 ON 20 MAY 2004  
CA INDEXING COPYRIGHT (C) 2004 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'USPAT2' ENTERED AT 11:31:02 ON 20 MAY 2004  
CA INDEXING COPYRIGHT (C) 2004 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'VETB' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'VETU' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'WPIDS' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'WPIFV' ENTERED AT 11:31:02 ON 20 MAY 2004  
COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'WPINDEX' ACCESS NOT AUTHORIZED

=> s MEGF10  
53 FILES SEARCHED...  
L1 16 MEGF10

=> DUP REM L1  
DUPLICATE IS NOT AVAILABLE IN 'ADISINSIGHT, ADISNEWS, BIOCOMMERCE, DGENE,  
DRUGMONOG2, IMSRESEARCH, FEDRIP, FOREGE, GENBANK, IMSPRODUCT, KOSMET,  
MEDICONF, NUTRACEUT, PCTGEN, PHAR, PHARMAML, RDISCLOSURE, SYNTHLINE'.  
ANSWERS FROM THESE FILES WILL BE CONSIDERED UNIQUE  
PROCESSING COMPLETED FOR L1  
L2 13DUP REM L1 (3 DUPLICATES REMOVED)

=> D L2 1-13

L2 ANSWER 1 OF 13 USPATFULL on STN  
AN 2004:7325 USPATFULL  
TI Proteins, polynucleotides encoding them and methods of using the same  
IN Anderson, David W., Branford, CT, UNITED STATES  
Burgess, Catherine E., Wethersfield, CT, UNITED STATES  
Caskan, Stacie J., North Haven, CT, UNITED STATES  
Colman, Steven D., Guilford, CT, UNITED STATES  
Edinger, Shlomit R., New Haven, CT, UNITED STATES  
Ellerman, Karen, Branford, CT, UNITED STATES  
Gerlach, Valerie, Branford, CT, UNITED STATES  
Gunther, Erik, Branford, CT, UNITED STATES  
Kekuda, Ramesh, Stamford, CT, UNITED STATES  
MacDougall, John R., Hamden, CT, UNITED STATES  
Mehraban, Fuad, Trumbull, CT, UNITED STATES  
Patturajan, Meera, Branford, CT, UNITED STATES  
Rothenberg, Mark, Clinton, CT, UNITED STATES  
Shimkets, Richard A., Guilford, CT, UNITED STATES  
Smithson, Glenda, Guilford, CT, UNITED STATES  
Spytek, Kimberly A., New Haven, CT, UNITED STATES  
Stone, David J., Guilford, CT, UNITED STATES  
Vernet, Corine A.M., Branford, CT, UNITED STATES  
Zerhusen, Bryan D., Branford, CT, UNITED STATES  
PI US 2004005558 A1 20040108  
AI US 2002-52648 A1 20020118 (10)  
PRAI US 2001-262454P 20010118 (60)  
US 2001-272920P 20010302 (60)  
US 2001-284549P 20010418 (60)  
US 2001-303229P 20010705 (60)  
US 2001-262892P 20010119 (60)  
US 2001-263605P 20010123 (60)  
US 2001-269098P 20010215 (60)  
US 2001-264159P 20010125 (60)  
US 2001-265517P 20010131 (60)  
US 2001-271855P 20010227 (60)  
US 2001-267057P 20010207 (60)  
US 2001-286287P 20010425 (60)  
DT Utility  
FS APPLICATION  
LN.CNT 10349  
INCL INCLM: 435/006.000  
INCLS: 435/007.200; 435/069.100; 435/320.100; 435/325.000; 514/012.000;  
514/044.000; 530/350.000; 530/388.100  
NCL NCLM: 435/006.000  
NCLS: 435/007.200; 435/069.100; 435/320.100; 435/325.000; 514/012.000;  
514/044.000; 530/350.000; 530/388.100  
IC [7]  
ICM: C12Q001-68  
ICS: G01N033-53; G01N033-567; A61K038-17; A61K048-00; C12P021-02;  
C12N005-06; C07K014-47; C07K016-18  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 2 OF 13 CAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 1  
AN 2001:330391 CAPLUS  
DN 136:64933  
TI Prediction of the coding sequences of unidentified human genes. XX. The  
complete sequences of 100 new cDNA clones from brain which code for large  
proteins in vitro  
AU Nagase, Takahiro; Nakayama, Manabu; Nakajima, Daisuke; Kikuno, Reiko;  
Ohara, Osamu  
CS Kazusa DNA Research Institute, Chiba, 292-0812, Japan  
SO DNA Research (2001), 8(2), 85-95  
CODEN: DARSE8; ISSN: 1340-2838  
PB Universal Academy Press  
DT Journal  
LA English  
RE.CNT 28 THERE ARE 28 CITED REFERENCES AVAILABLE FOR THIS RECORD  
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 3 OF 13 LIFESCI COPYRIGHT 2004 CSA on STN  
AN 2004:35501 LIFESCI  
TI Prediction of the Coding Sequences of Unidentified Human Genes. XX. The  
Complete Sequences of 100 New cDNA Clones from Brain which Code for Large  
Proteins in vitro  
AU Nagase, T.; Nakayama, M.; Nakajima, D.; Kikuno, R.; Ohara, O.

CS Kazusa DNA Research Institute, 1532-3 Yana, Kisarazu, Chiba 292-0812,  
Japan  
SO DNA Research [DNA Res.], (20010000) vol. 8, [np].  
ISSN: 1340-2838.  
DT Journal  
FS N  
LA English  
SL English

L2 ANSWER 4 OF 13 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN  
AN AAL52376 DNA DGENE  
TI New epidermal growth factor producing Lactococcus lactis or Lactococcus  
casei, useful for promoting gut adsorption or for treating Short Bowel  
Syndrome.  
IN Steidler L  
PA (VLAA-N) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
(UYGE-N) UNIV GENT.  
PI WO 2004001020 A2 20031231  
AI WO 2003-EP50242 20030619  
PRAI EP 2002-77532 20020619  
DT Patent  
LA English  
OS 2004-071734 [07]  
DESC Murine epidermal growth factor coding sequence oligo \*\*\*MEGF10\*\*\* .

L2 ANSWER 5 OF 13 GENBANK.RTM. COPYRIGHT 2004 on STN

LOCUS (LOC): CK656927 GenBank (R)  
GenBank ACC. NO. (GBN): CK656927  
GenBank VERSION (VER): CK656927.1 GI:44893777  
CAS REGISTRY NO. (RN): 660084-00-4  
SEQUENCE LENGTH (SQL): 483  
MOLECULE TYPE (CI): mRNA; linear  
DIVISION CODE (CI): Expressed sequence tag  
DATE (DATE): 3 Mar 2004  
DEFINITION (DEF): 3017HFEX25F2 BgORESTES schistosoma-exposed NHM 3017  
Head/foot 2 Biomphalaria glabrata cDNA clone ZBA0961  
similar to \*\*\*MEGF10\*\*\* protein, mRNA sequence.  
KEYWORDS (ST): EST  
SOURCE: Biomphalaria glabrata (bloodfluke planorb)  
ORGANISM (ORGN): Biomphalaria glabrata  
Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata;  
Basommatophora; Lymnaeoidea; Planorbidae; Biomphalaria

COMMENT:  
Contact: Lockyer, A.E.  
Wolfson Wellcome Biomedical Laboratory  
The Natural History Museum  
Cromwell Road, London, SW7 5BD, UK  
Tel: +44 (0)20 7942 5148  
Fax: +44 (0)20 7942 5518  
Email: a.lockyer@nhm.ac.uk  
High quality sequence stop: 483  
POLYA=No.

REFERENCE: 1 (bases 1 to 483)  
AUTHOR (AU): Lockyer,A.E.; Spinks,J.N.; Kane,R.A.; Dias Neto,E.;  
Noble,L.R.; Rollinson,D.; Jones,C.S.  
TITLE (TI): ESTs from Biomphalaria glabrata using the ORESTES  
approach  
JOURNAL (SO): Unpublished (2003)

#### FEATURES (FEAT):

| Feature Key | Location | Qualifier  |
|-------------|----------|--|
| source      | 1..483   | /organism="Biomphalaria glabrata"<br>/mol-type="mRNA"<br>/strain="NHM 3017"<br>/db-xref="taxon:6526"<br>/clone="ZBA0961"<br>/sex="Hermaphrodite"<br>/tissue-type="Head/foot"<br>/clone-lib="BgORESTES<br>schistosoma-exposed NHM 3017<br>Head/foot 2"<br>/note="Vector: pGEM; mRNA,<br>extracted from schistosoma-exposed<br>3017 (resistant) snails was used" |

as a template for RT-PCR with random primers to generate cDNA fragments. These were cloned and sequenced using M13F. Primer sequences were removed from sequence."

SEQUENCE (SEQ):

```
1 ggtcagaact gttccaatgt ctgccaatgt gatgtcagta attctctaaa ctgtaatgat
61 gtcgatggta catgcacttg taaaactggc tggactggaa ccaactgtga acaggacata
121 gatgagtgtg ctattaattt aacattctgc tcaaactctt atgaagtttg ccgcaatgtg
181 aagggctctg cggaatgtat ttgtcaagat gggtttctaca attcaacaat tttattcagt
241 tgtcaagcct gtgactctct ccactatgga tcaaactgca cctctgtatg ctcatgtcat
301 acaaccaaca cagcagattg taatgatgtc aatggaacat gttcttgtaa acctggatgg
361 actggggctg attgctcgca aggttgatgt tccctgcact atgggtcaaga ctgcaccttg
421 caatgtaa atgtcagttaaa taacagcgct gtatgcaata gcactgatgg aactgcaca
481 tgt
```

L2 ANSWER 6 OF 13 GENBANK.RTM. COPYRIGHT 2004 on STN

LOCUS (LOC): CD648703 GenBank (R)  
GenBank ACC. NO. (GBN): CD648703  
GenBank VERSION (VER): CD648703.1 GI:31905003  
CAS REGISTRY NO. (RN): 537522-34-2  
SEQUENCE LENGTH (SQL): 801  
MOLECULE TYPE (CI): mRNA; linear  
DIVISION CODE (CI): Expressed sequence tag  
DATE (DATE): 18 Jun 2003  
DEFINITION (DEF): AUF\_103\_I07\_T7 *Crassostrea virginica* Gill *Crassostrea virginica* cDNA 5' similar to \*\*\*MEGF10\*\*\* protein, mRNA sequence.

KEYWORDS (ST): EST  
SOURCE: *Crassostrea virginica* (eastern oyster)  
ORGANISM (ORGN): *Crassostrea virginica*  
Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia;  
Ostreoida; Ostreoidae; Ostreidae; *Crassostrea*

NUCLEIC ACID COUNT (NA): 253 a 130 c 157 g 261 t

COMMENT:

Contact: Liu ZJ  
The Fish Molecular Genetics and Biotechnology Laboratory,  
Department of Fisheries and Allied Aquacultures and Program of Cell  
and Molecular Biosciences  
Auburn University  
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA  
Tel: 334 844 4054  
Fax: 334 844 9208  
Email: zliu@acesag.auburn.edu  
Seq primer: M13 Reverse.

REFERENCE: 1 (bases 1 to 801)  
AUTHOR (AU): Peatman,E.; Kucuktas,H.; Li,P.; He,C.; Feng,J.; Wei,X.;  
Liu,Z.  
TITLE (TI): Differentially expressed oyster (*Crassostrea virginica*)  
genes after exposure to mercury  
JOURNAL (SO): Unpublished

FEATURES (FEAT):

| Feature Key | Location | Qualifier  |
|-------------|----------|--|
| source      | 1..801   | /organism="Crassostrea virginica"<br>/mol-type="mRNA"<br>/db-xref="taxon:6565"<br>/clone-lib="Crassostrea virginica<br>Gill"<br>/note="Organ: Gill; Vector:<br>pSport1; Site-1: NotI; Site-2:<br>SalI" |

SEQUENCE (SEQ):

```
1 gtttttgctt agaagcggat gggtaatcgg gtacattttc tgaactgggtg gtgatagggtg
61 ccttctgtag atgtgtgatg gcatataaac gaattcgatc acatatTTTT ttcgtcattt
121 tcaatctcct ataattgaaa tgtattgttt aacgccgaaa catcatgatt gagcccattt
181 agccccgtgt gttggtctac ctttattcca atcataaaat cgatacttgt agaatgcagt
241 aatggatatt acggttaaca ctgtaatat acgtgcgggt attgcctaca aggaagtgtg
301 tgcaatagac tgaatggaac ttgtgaaaat ggggtgatca accattttta agaaccgaga
361 tgtgctgttt gtagagacgg attctacaac agtagatgca cttctcaatg cggcaagtgt
421 gtgaacgatg agccttgtga taaagtaaca ggagaatgca ggaatggatg tcaacaacat
```

481 ttggaacctc ctttgtgtca gggtttgtaa tgcataatat attgtgtact gatatgcacg  
541 cttcaacgtg ttatatgata tattgcattc aaactgtctc atagcattga ccagcattgt  
601 aaagtcaaaa ctgcgttttg cattattaat gtatttcacg taatgaattt cctgtataca  
661 gttatgttaa cgtcttaaaa acacgcttta aactacacgt ccaaaataca tttatttgtg  
721 acctgtgtaa gaatatagat aaaataagta ggtttcttct taaccattaa aaacaaaaca  
781 atacaaaatc gtgtttaaac a

L2 ANSWER 7 OF 13 GENBANK.RTM. COPYRIGHT 2004 on STN

LOCUS (LOC): CD646852 GenBank (R)  
GenBank ACC. NO. (GBN): CD646852  
GenBank VERSION (VER): CD646852.1 GI:31901304  
CAS REGISTRY NO. (RN): 537503-80-3  
SEQUENCE LENGTH (SQL): 572  
MOLECULE TYPE (CI): mRNA; linear  
DIVISION CODE (CI): Expressed sequence tag  
DATE (DATE): 18 Jun 2003  
DEFINITION (DEF): AUF\_106\_F02\_T7 Crassostrea virginica Gonad Crassostrea virginica cDNA 5' similar to \*\*\*MEGF10\*\*\* protein, mRNA sequence.  
KEYWORDS (ST): EST  
SOURCE: Crassostrea virginica (eastern oyster)  
ORGANISM (ORGN): Crassostrea virginica  
Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Ostreoida; Ostreioidea; Ostreidae; Crassostrea  
NUCLEIC ACID COUNT (NA): 174 a 85 c 144 g 169 t  
COMMENT:

Contact: Liu ZJ  
The Fish Molecular Genetics and Biotechnology Laboratory,  
Department of Fisheries and Allied Aquacultures and Program of Cell  
and Molecular Biosciences  
Auburn University  
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA  
Tel: 334 844 4054  
Fax: 334 844 9208  
Email: zliu@acesag.auburn.edu  
Seq primer: M13 Reverse.

REFERENCE: 1 (bases 1 to 572)  
AUTHOR (AU): Peatman,E.; Kucuktas,H.; Li,P.; He,C.; Feng,J.; Wei,X.; Liu,Z.  
TITLE (TI): Differentially expressed oyster (Crassostrea virginica) genes after exposure to mercury  
JOURNAL (SO): Unpublished

FEATURES (FEAT):  

| Feature Key | Location | Qualifier   |
|-------------|----------|---|
| source      | 1..572   | /organism="Crassostrea virginica"<br>/mol-type="mRNA"<br>/db-xref="taxon:6565"<br>/clone-lib="Crassostrea virginica Gonad"<br>/note="Organ: Gonad; Vector: pSport1; Site-1: NotI; Site-2: SalI" |

SEQUENCE (SEQ):  
1 gtgatagtg atgtaaaccg ggctggaagg gacaatactg tcacgaacaa tgtcaatatg  
61 gcttctatgg acaggactgc ctacttacat gtaacgacaa atgtaacggg tgaacaatg  
121 tgaacgggta ctgtgataga ggatgtaaac caggctggaa gggggacaac tgtcaacaac  
181 aatgcatgtc tgacttttat ggcgaaaact gtagcagaaa ttgtggatat tgcttagatg  
241 gtagaacatg ccaccacata aatggcacat gtgaacaggg atgtaatccg ggttacaag  
301 ctccattctg taacgaagag tgtgattttg gattctacgg aaatagatgc cttcaggaat  
361 gtgggttcatt ttgcaaaaata tcacgtgatt gtcattcatt gactggtttt tgtaaaaacg  
421 gctgtaaaag tggttggcaa ggaaaggatt gttttgaagt ttcgaaactt gaagacagta  
481 atacagattg gaaatcaaga ttttacggaa tgttgggggc gttcttcgta ttgttgattt  
541 tgatttgtct tcttattgcc taccatatta ta

L2 ANSWER 8 OF 13 GENBANK.RTM. COPYRIGHT 2004 on STN

LOCUS (LOC): CD034532 GenBank (R)  
GenBank ACC. NO. (GBN): CD034532  
GenBank VERSION (VER): CD034532.1 GI:30416370  
CAS REGISTRY NO. (RN): 514024-90-9  
SEQUENCE LENGTH (SQL): 471  
MOLECULE TYPE (CI): mRNA; linear



DIVISION CODE (CI): Expressed sequence tag  
DATE (DATE): 7 May 2003  
DEFINITION (DEF): mgmt016xK13f.b Mated culture Magnaporthe grisea cDNA  
clone mgmt016xK13 5', mRNA sequence.  
KEYWORDS (ST): EST  
SOURCE: Magnaporthe grisea (anamorph: Pyricularia grisea)  
ORGANISM (ORGN): Magnaporthe grisea  
Eukaryota; Fungi; Ascomycota; Pezizomycotina;  
Sordariomycetes; Sordariomycetes incertae sedis;  
Magnaporthaceae; Magnaporthe  
NUCLEIC ACID COUNT (NA): 116 a 125 c 91 g 139 t  
COMMENT:

Contact: Ebbole DJ  
Department of Plant Pathology & Microbiology  
Texas A&M University  
Peterson Bldg, MS2132, College Station, TX 77843-2132, USA  
Tel: 979 845 4831  
Fax: 979 845 6483  
Email: d-ebbole@tamu.edu  
Chromatogram file of this sequence is available, see contact person  
;Best nr hit (April. 22, 2003) ref|XP\_236337.1| similar to \*\*\*MEGF10\*\*\*  
protein [Homo sapiens] [Rattu... 33 1.2  
PCR Primers  
FORWARD: T3 primer  
BACKWARD: T7 primer  
Plate: mgmt016 row: K column: 13  
Seq primer: T3.  
REFERENCE: 1 (bases 1 to 471)  
AUTHOR (AU): Ebbole,D.J.; Yuan,J.; Thomas,T.L.; Bobrowicz,P.; Lu,G.;  
Bhatterai,K. ; Dean,R.A.  
TITLE (TI): Expressed sequence tags from the rice blast fungus,  
Magnaporthe grisea  
JOURNAL (SO): Unpublished (2002)

| FEATURES (FEAT): |          |   |
|------------------|----------|---|
| Feature Key      | Location | Qualifier   |
| source           | 1..471   | /organism="Magnaporthe grisea"<br>/strain="4091-5-8 X 4136-4-3"<br>/db-xref="taxon:148305"<br>/clone="mgmt016xK13"<br>/clone-lib="Mated culture"<br>/sex="Mat1-2 and Mat1-1 mixed culture"<br>/cell-type="mixed sexual development"<br>/dev-stage="asci, ascospores, perithecia, mycelium"<br>/note="Vector: pBluescriptSK; Site-1: EcoRI; Site-2: XhoI; Two mating types were co-cultivated over a filter paper on oatmeal agar medium. After three days at 25 C plates were transferred to 21 C. Perithecia with asci and ascospores formed at the beginning of the third week. Material was collected by scraping tissue from the filter paper. Sequences were processed by one of two methods. where a full-length alignment to the M. grisea genome sequence was available, the EST sequence was trimmed according to the alignment, otherwise sequence quality was assessed using phredPhrap version 991019 and trimmed according to phd files (0.05) and for vector seqs." |

SEQUENCE (SEQ):  
1 ggcttttctt tctttttcgt ttgaccagac ccaagccgca aatccatcca acgatatact  
61 ttttcttctt tttttttctt ctcttttctt cgtaactct caccgaatta ctgccgacat  
121 aggccttcta gtactagga cttctcttga aagcaggtag atcattctcg acccgaaacg  
181 ccgaaaataa aatcgggatt tccgcgatcc gcgtttcgca ttttcaacga caaaccatta  
241 ccaagaggag gactcggcct tatcaaaacc atcgcgctca tcctagaagt gctgctgttt

301 tctactacac caagaacact gcctcgctga atttatttagc gaacgaacga agatttctctg  
 361 gcatgtcggg catcacacag gcaacggcca tggtggccaa gcgagctgat ggatgtcccg  
 421 atggatacta ccggtctgga aactactgtt atcggggatc tggatggtac t

L2 ANSWER 9 OF 13 GENBANK.RTM. COPYRIGHT 2004 on STN

LOCUS (LOC): BU493977 GenBank (R)  
 GenBank ACC. NO. (GBN): BU493977  
 GenBank VERSION (VER): BU493977.1 GI:22790235  
 CAS REGISTRY NO. (RN): 453296-65-6  
 SEQUENCE LENGTH (SQL): 257  
 MOLECULE TYPE (CI): mRNA; linear  
 DIVISION CODE (CI): Expressed sequence tag  
 DATE (DATE): 1 Oct 2002  
 DEFINITION (DEF): vaa03f07.y2 SV DOGRDA RD3 Canis familiaris cDNA 5' similar to gp|AAH20198.1|BC020198\_1 similar to \*\*\*MEGF10\*\*\* protein [Homo ;, mRNA sequence.  
 SOURCE: dog.  
 ORGANISM (ORGN): Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis  
 NUCLEIC ACID COUNT (NA): 67 a 65 c 60 g 65 t  
 COMMENT:  
 other\_ESTs: vaa03f07.x2  
 Contact: Skip Virgin  
 RNA Expression in Diseased Tissues by RDA  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Library was constructed by Tianxiang Shen and Skip Virgin DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Skip Virgin (virgin@immunology.wustl.edu)  
 Possible reversed clone: similarity on wrong strand  
 Seq primer: -40RP from Gibco.  
 REFERENCE: 1 (bases 1 to 257)  
 AUTHOR (AU): Virgin,S.; Tianxiang,S.; Karst,S.; Wobus,C.; Lay,M.; Clifton,S.; Pape,D.; Marra,M.; Hillier,L.; Martin,J.; Wylie,T.; Dante,M.; Theising,B.; Bowers,Y.; Gibbons,M.; Ritter,E.; Bennett,J.; Ronko,I.; Tsagareishvili,R.; Maguire,L.; Kennedy,S.; Waterston,R.; Wilson,R.  
 TITLE (TI): RNA Expression in Diseased Tissues by RDA  
 JOURNAL (SO): Unpublished (2002)

FEATURES (FEAT):

| Feature Key | Location | Qualifier   |
|-------------|----------|---|
| source      | 1..257   | /organism="Canis familiaris"<br>/db-xref="taxon:9615"<br>/clone-lib="SV DOGRDA RD3"<br>/tissue-type="Brain from a dog with encephalitis"<br>/lab-host="DH5a"<br>/note="Vector: pCR2.1-TOPO; PCR products were directly cloned into TA cloning vector (pre-cut vector pCR2.1-TOPO was purchased from Invitrogen) Average insert size is 300bp. Our Insert also have same adaptor sequence on both ends. The sequence of the adaptor is: 5'-ATACGTGCAGGCTGGTTACATCTG-3'." |

SEQUENCE (SEQ):  
 1 cacaggcact ggagcagttg gttccacccc agccaggctc aactgacag gtgtttggag  
 61 caatgcagcg accatggaca catttatcag cacagtgggg gacacacatt tccctgcttt  
 121 cataaaatcc aggacaacac tgggatttgc gcctatacat agttttctcc ccatgtcgtat  
 181 aggcagtccg ataactgatt ctgtgccgtg tacatttaaa ccagttcaga atatcagtg  
 241 agctggtgta gtagatt

L2 ANSWER 10 OF 13 GENBANK.RTM. COPYRIGHT 2004 on STN

LOCUS (LOC): BU493770 GenBank (R)

GenBank ACC. NO. (GBN): BU493770  
 GenBank VERSION (VER): BU493770.1 GI:22790028  
 CAS REGISTRY NO. (RN): 453296-12-3  
 SEQUENCE LENGTH (SQL): 311  
 MOLECULE TYPE (CI): mRNA; linear  
 DIVISION CODE (CI): Expressed sequence tag  
 DATE (DATE): 1 oct 2002  
 DEFINITION (DEF): vaa03f07.x2 SV DOGRDA RD3 Canis familiaris cDNA 3' similar to gp|BAB47409.1|AB058676\_1 \*\*\*MEGF10\*\*\* protein ;, mRNA sequence.  
 SOURCE: dog.  
 ORGANISM (ORGN): Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis  
 NUCLEIC ACID COUNT (NA): 82 a 71 c 75 g 82 t 1 others  
 COMMENT:  
 other\_ESTs: vaa03f07.y2  
 Contact: Skip Virgin  
 RNA Expression in Diseased Tissues by RDA  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Library was constructed by Tianxiang Shen and Skip Virgin DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Skip Virgin (virgin@immunology.wustl.edu)  
 Putative full length read  
 vector to vector length is 312  
 Possible reversed clone: similarity on wrong strand  
 Seq primer: -40UP from Gibco.  
 REFERENCE: 1 (bases 1 to 311)  
 AUTHOR (AU): Virgin,S.; Tianxiang,S.; Karst,S.; Wobus,C.; Lay,M.; Clifton,S.; Pape,D.; Marra,M.; Hillier,L.; Martin,J.; Wylie,T.; Dante,M.; Theising,B.; Bowers,Y.; Gibbons,M.; Ritter,E.; Bennett,J.; Ronko,I.; Tsagareishvili,R.; Maguire,L.; Kennedy,S.; Waterston,R.; Wilson,R.  
 TITLE (TI): RNA Expression in Diseased Tissues by RDA  
 JOURNAL (SO): Unpublished (2002)

| Feature Key | Location | Qualifier   |
|-------------|----------|---|
| source      | 1..311   | /organism="Canis familiaris"<br>/db-xref="taxon:9615"<br>/clone-lib="SV DOGRDA RD3"<br>/tissue-type="Brain from a dog with encephalitis"<br>/lab-host="DH5a"<br>/note="Vector: pCR2.1-TOPO; PCR products were directly cloned into TA cloning vector (pre-cut vector pCR2.1-TOPO was purchased from Invitrogen) Average insert size is 300bp. Our Insert also have same adaptor sequence on both ends. The sequence of the adaptor is: 5'-ATACGTGCAGGCTGGTTACATCTG-3'." |

SEQUENCE (SEQ):  
 1 tgatcaaatc tactacacca gctgcactga tattctgaac tggtttaa at gtacacggca  
 61 cagaatcagt tatcggactg cctatcgaca tggggagaaa actatgtata ggcgcaaatc  
 121 ccagtgttgt cctggatttt atgaaagcag ggaaatgtgt gtccccact gtgctgataa  
 181 atgtgtccat ggtcgtcgca ttgctccaaa cacctgtcag tgtgagcctg gctgggggtg  
 241 aaccaactgc tccagtgctt gtgatgggtg tcagatgtaa ccagcctgca cgtataatca  
 301 ctantgaatt c

L2 ANSWER 11 OF 13 GENBANK.RTM. COPYRIGHT 2004 on STN

LOCUS (LOC): BM427205 GenBank (R)  
 GenBank ACC. NO. (GBN): BM427205  
 GenBank VERSION (VER): BM427205.1 GI:18432195  
 CAS REGISTRY NO. (RN): 396916-69-1  
 SEQUENCE LENGTH (SQL): 668

MOLECULE TYPE (CI): mRNA; linear  
DIVISION CODE (CI): Expressed sequence tag  
DATE (DATE): 30 Jan 2002  
DEFINITION (DEF): pgf2n.pk006.a15 Normalized Chicken Abdominal Fat Library (pgf2n) Gallus gallus cDNA clone  
pgf2n.pk006.a15 5' similar to gi|14192943  
ref|NP\_115822.1| \*\*\*MEGF10\*\*\* protein [Homo sapiens] gi|14724016 ref|XP\_030163.1| \*\*\*MEGF10\*\*\* protein [Homo sapiens] dbj|BAB47409.1| (AB058676) \*\*\*MEGF10\*\*\* protein (KIAA1780) [Homo sapiens], mRNA sequence.

SOURCE: chicken.  
ORGANISM (ORGN): Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus

NUCLEIC ACID COUNT (NA): 108 a 256 c 199 g 104 t 1 others  
COMMENT:

Contact: Larry A. Coghurn  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1335  
Fax: 302-831-2822  
Email: coghurn@udel.edu, www.chickest.udel.edu.

REFERENCE: 1 (bases 1 to 668)  
AUTHOR (AU): Coghurn, L.A.; Morgan, R.; Burnside, J.  
TITLE (TI): ESTs from Normalized Chicken fat cDNA library-USDA/IFAFS Animal Genome Project  
JOURNAL (SO): Unpublished (2002)

| FEATURES (FEAT): |          |  |
|------------------|----------|--|
| Feature Key      | Location | Qualifier  |
| source           | 1..668   | /organism="Gallus gallus"<br>/strain="Commercial broiler, Ottawa Research Centre strains, commercial leghorn"<br>/db-xref="taxon:9031"<br>/clone="pgf2n.pk006.a15"<br>/clone-lib="Normalized Chicken Abdominal Fat Library (pgf2n)"<br>/sex="Male and Female"<br>/tissue-type="Abdominal Fat"<br>/dev-stage="Embryonic (d18,d19); post-hatch (d1,w3,w7,w9 ,w16,1yr)"<br>/lab-host="E. coli EMDH10B"<br>/note="Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each developmental age (across strains); single pass sequencing from 5'-end" |

SEQUENCE (SEQ):  
1 ccacccacgc gaccccaacg tctgcagcta ctgggagagt ttacagtag cagtgaagga  
61 gtcatacgcc aaacccca ca ttgtgacctc caccgagccc tgcactggga caccaggatt  
121 gccacagtcg tgcccgcagc aaagggtcgt gtaccgcacg gactaccgac aggcggtgcg  
181 caccgactac cgccggcgct accagtgtct cctgggctac tatgagagca gggacgcctg  
241 cgtcccgcgc tgcacccagg agtgcgtcca cgggcggtgc gtggcccccg agcgttgcca  
301 gtgtgagccg ggctggaggg gacacgactg ctccagcgcg tgtgatgagc gtttgtgggg  
361 caagaactgc gagcaccact gtgactgcca ccacggggcc ccctgcaatc ccctgagtgg  
421 ggtctgtgcc tgccccctcg gcttcacccc cccactgtgc caccagccct gcccgcccg  
481 ctccctacgg cccgactgcc gctctgactg cccatgccac cacggggccc cctgcaacgc  
541 ctccactggg gcctgcctct gccccccagg acttgctggc ccactctgtg aggtgctgtg  
601 ccccgagggg acaccgtgcg gcacccactg cccctgccag aatgggggta tctgccaccc  
661 ccnngcct

LOCUS (LOC): BC020198 GenBank (R)  
GenBank ACC. NO. (GBN): BC020198  
GenBank VERSION (VER): BC020198.1 GI:18044365  
CAS REGISTRY NO. (RN): 381931-64-2  
SEQUENCE LENGTH (SQL): 2267  
MOLECULE TYPE (CI): mRNA; linear  
DIVISION CODE (CI): Primates

DATE (DATE): 16 Sep 2003  
 DEFINITION (DEF): Homo sapiens \*\*\*MEGF10\*\*\* protein, mRNA (cDNA clone  
 IMAGE:4904255), complete cds.  
 SOURCE: Homo sapiens (human)  
 ORGANISM (ORGN): Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
 Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini;  
 Hominidae; Homo  
 NUCLEIC ACID COUNT (NA): 543 a 548 c 631 g 545 t  
 COMMENT:

Contact: MGC help desk  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bcgsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline  
 Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,  
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 40 Row: o Column: 4  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 14192942  
 This clone has the following problem: The cds is short compared to  
 the longest cds in the locus.

REFERENCE: 1 (bases 1 to 2267)  
 AUTHOR (AU): Strausberg,R.L.; Feingold,E.A.; Grouse,L.H.;  
 Derge,J.G.; Klausner,R.D.; Collins,F.S.; Wagner,L.;  
 Shenmen,C.M.; Schuler,G.D.; Altschul,S.F.; Zeeberg,B.;  
 Buetow,K.H.; Schaefer,C.F.; Bhat,N.K.; Hopkins,R.F.;  
 Jordan,H.; Moore,T.; Max,S.I.; Wang,J.; Hsieh,F.;  
 Diatchenko,L.; Marusina,K.; Farmer,A.A.; Rubin,G.M.;  
 Hong,L.; Stapleton,M.; Soares,M.B.; Bonaldo,M.F.;  
 Casavant,T.L.; Scheetz,T.E.; Brownstein,M.J.;  
 Usdin,T.B.; Toshiyuki,S.; Carninci,P.; Prange,C.;  
 Raha,S.S.; Loquellano,N.A.; Peters,G.J.; Abramson,R.D.;  
 Mullahy,S.J.; Bosak,S.A.; McEwan,P.J.; McKernan,K.J.;  
 Malek,J.A.; Gunaratne,P.H.; Richards,S.; Worley,K.C.;  
 Hale,S.; Garcia,A.M.; Gay,L.J.; Hulyk,S.W.;  
 Villalon,D.K.; Muzny,D.M.; Sodergren,E.J.; Lu,X.;  
 Gibbs,R.A.; Fahey,J.; Helton,E.; Kettelman,M.; Madan,A.;  
 Rodrigues,S.; Sanchez,A.; Whiting,M.; Madan,A.;  
 Young,A.C.; Shevchenko,Y.; Bouffard,G.G.;  
 Blakesley,R.W.; Touchman,J.W.; Green,E.D.;  
 Dickson,M.C.; Rodriguez,A.C.; Grimwood,J.; Schmutz,J.;  
 Myers,R.M.; Butterfield,Y.S.; Krzywinski,M.I.;  
 Skalska,U.; Smailus,D.E.; Schnerch,A.; Schein,J.E.;  
 Jones,S.J.; Marra,M.A.

TITLE (TI): Generation and initial analysis of more than 15,000  
 full-length human and mouse cDNA sequences  
 JOURNAL (SO): Proc. Natl. Acad. Sci. U.S.A., 99 (26), 16899-16903  
 (2002)

OTHER SOURCE (OS): CA 138:84323  
 REFERENCE: 2 (bases 1 to 2267)  
 AUTHOR (AU): Strausberg,R.  
 TITLE (TI): Direct Submission  
 JOURNAL (SO): Submitted (19-DEC-2001) National Institutes of Health,  
 Mammalian Gene Collection (MGC), Cancer Genomics  
 Office, National Cancer Institute, 31 Center Drive,  
 Room 11A03, Bethesda, MD 20892-2590, USA

# FEATURES (FEAT):

| Feature Key | Location | Qualifier   |
|-------------|----------|---|
| source      | 1..2267  | /organism="Homo sapiens"<br>/mol-type="mRNA"<br>/db-xref="taxon:9606"<br>/clone="IMAGE:4904255" |

```

/tissue-type="Muscle,
rhabdomyosarcoma"
/clone-lib="NIH-MGC-17"
/lab-host="DH10B-R"
/note="Vector: pOTB7"
/gene="MEGF10"
/note="synonym: KIAA1780"
/db-xref="LocusID:84466"
/codon-start=1
/product="MEGF10 protein"
/protein-id="AAH20198.1"
/db-xref="GI:18044366"
/db-xref="LocusID:84466"
/translation="MVISLNSCLSFICLLCHWI
GTASPLNLEDPNVCSHWESYSVTV
QESYPHPFDQIYYTSCTDILNWFKCTRHRVSYRT
AYRHGEKTMYYRRKSQCCPGFYESG
EMCVPHCADKCVHGRCIAPNTCQCEPGWGGTNCS
SACDGDHWGPHCTSRCQCKNGALC
NPITGACHCAAGFRGWRCEDRCEQGTYGNDCHQR
CQCQNGATCDHVTGECRCPPGYTG
AFCEDLCPPGKHGPQCEQRCPCQNGGVCHHVTGE
CSCPSGWMGTVCGQPCPEGRFGKN
CSQECQCHNGGTCDAAATGQCHCSPGYTGERCQDE
CPVGTYGVLCAETCQCVNGGKCYH
VSGACLCEAGFAGERCEARLCPEGLYGIKCDKRC
PCHLENTHSCHPMSGECACKPGWS
GLYCNETCSPGFYGEACQICSCQNGADCDSVTG
KCTCAPGFGIDCSTPCPLGTYGI
NCSSRCGCKNDAVCSVDGSCCKAGWHGVDCSI
RCPSTGWGFGCNLTQCLNGGACN
TLDGTCTCAPGWRGEKCELPQDGTYGLNCAERC
DCSHADGCHPTTGHCRLPGWSGL F"

```

gene 1..2267

CDS 280..1983

SEQUENCE (SEQ):

```

1 aaactgggagc actgggaacg cgttgagacg ttcctctttc ccgcttctcc acctttacgc
61 ctgaaagaag actcccaaag attgctttct tctgggacgc tgcttgaacg ctaaccgcgt
121 tgattggaac agattttgtg tcttggtgg ctttgggtga agaccgggga gaaaagggtt
181 cgctgcgatt ctcaagatct ctggacctgg attatcattg caaatccatt aaagaaggag
241 aagcaagcgg atttcagaga ggttggttctt cagaaaaaaa tggttatttc tttgaactca
301 tgcctgagct ttatttggtt attgttatgc cactggattg ggacagcatc acctctgaat
361 cttgaagacc ctaatgtgtg tagccactgg gaaagctact cagtgaactgt gcaagagtca
421 taccacatc cttttgatca aattttactac acgagctgca ctgacattct aaactggttt
481 aaatgcacgc ggcacagagt cagctatcgg acagcctatc gacatgggga gaagactatg
541 tataggcgca agtctcagtg ttgtcctgga ttttatgaaa gcggggaaat gtgtgtcccc
601 cactgtgctg ataaatgtgt ccatggtcgc tgtattgctc caaacacctg tcagtgtgag
661 cctggctggg gagggaccaaa ctgctccagt gcctgcgatg gtgatcactg ggggtccccac
721 tgaccagacc ggtgccagtg caaaaatggg gctctgtgca accccatcac cggggcttgc
781 cactgtgctg cgggcttccg gggctggcgc tgcgaggacc gctgtgagca gggcacctat
841 ggtaacgact gtcacacagag atgccagtgcc cagaatggag ccacctgcga ccacgtcacg
901 ggggaatgcc gctgcccacc aggatacacc ggagccttct gtgaggatct ttgtcctcct
961 ggtaaacatg gtccacagtg tgagcagaga tggccttctc aaaatggagg agtgtgtcat
1021 cacgtcactg gagaatgtct ttgccccttct ggctggatgg gcacagtgtg tggtcagcct
1081 tgccccgagg gtcgcttttg aaagaactgt tcccaagaat gccagtgcga taatggaggg
1141 acgtgtgatg ctgccacagg ccaatgtcat tgcagtccag gatacacagg ggaacgggtg
1201 caggatgagt gtcctgtttg gacctatggc gttctctgtg ctgagacctg ccagtgtgtc
1261 aacggaggga agtgtttacca cgtgagcggc gcatgcctct gtgaagcagg ctttgcctgg
1321 gagcgtgctg aagcacgcct gtgtcctgag gggctctacg gcatcaaatg tgacaaacgg
1381 tgtccctgcc acttggaaaa cactcatagc tgtcacccca tgtctggaga gtgtgcctgc
1441 aagccgggct ggtcaggact ctactgtaat gagacatgtt ctcttgatt ctacggggaa
1501 gcttgccagc agatctgcag ctgccaaaat ggggcagact gtgacagtgt gactggaaaag
1561 tgcacctgtg ccccaggatt caaaggaatt gactgtctta ccccatgccc tctgggaacc
1621 tatgggataa actgttccct tcgctgtggc tgtaaaaatg atgcagtctg ctctcctgtg
1681 gacgggtctt gtacttgcaa ggcagggtgg caccgggtgg actgctccat cagatgtccc
1741 agtggcacat ggggcttttg ctgtaactta acatgccagt gcctcaacgg gggagcctgc
1801 aacaccctgg acgggacctg cacgtgtgca cctggatggc gcggggagaa atgcgaactt
1861 ccctgccagg atggcacgta cgggctgaac tgtgctgagc gctgcgactg cagccacgca
1921 gatggctgac accctaccac gggccattgc cgctgcctcc ccgatgggtc aggattgttt
1981 tgactattct gggctcctga atttccatgt gaatggtaac atcaatatgt taatttccac
2041 aaagaagcaa ggtggaatgt taatggagat tgaatctata gatcaatttg ggaataactg
2101 tccacttaac aatatgaagt cttcagattc ataaacatgg gatgtcacca gcaacctctt
2161 cagtagtgct tctcagcccc ttggaaaatg ttgagagctt ctttggtttc taaaggagta
2221 aactgagatg aactgtcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa

```

LOCUS (LOC): AB058676 GenBank (R)  
GenBank ACC. NO. (GBN): AB058676  
GenBank VERSION (VER): AB058676.1 GI:14017776  
CAS REGISTRY NO. (RN): 336671-12-6  
SEQUENCE LENGTH (SQL): 7522  
MOLECULE TYPE (CI): mRNA; linear  
DIVISION CODE (CI): Primates  
DATE (DATE): 10 May 2001  
DEFINITION (DEF): Homo sapiens mRNA for \*\*\*MEGF10\*\*\* protein  
(KIAA1780), complete cds.  
SOURCE: Homo sapiens adult hippocampus cDNA to mRNA,  
clone:pf01012.  
ORGANISM (ORGN): Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini;  
Hominidae; Homo  
NUCLEIC ACID COUNT (NA): 2130 a 1554 c 1697 g 2141 t  
REFERENCE: 1 (sites)  
AUTHOR (AU): Nagase,T.; Nakayama,M.; Nakajima,D.; Kikuno,R.;  
Ohara,O.  
TITLE (TI): Prediction of the coding sequences of unidentified  
human genes. XX. The complete sequences of 100 new cDNA  
clones from brain which code for large proteins in  
vitro  
JOURNAL (SO): DNA Res., 8 (2), 85-95 (2001)  
OTHER SOURCE (OS): CA 136:64933  
REFERENCE: 2 (bases 1 to 7522)  
AUTHOR (AU): Nakayama,M.; Nagase,T.; Nakajima,D.; Kikuno,R.;  
Ohara,O.  
TITLE (TI): Direct Submission  
JOURNAL (SO): Submitted (27-MAR-2001) Manabu Nakayama, Kazusa DNA  
Research Institute, Department of Human Gene Research;  
1532-3, Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:nmanabu@kazusa.or.jp,  
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3915,  
Fax:81-438-52-3914)

| FEATURES (FEAT): |           |   |
|------------------|-----------|---|
| Feature Key      | Location  | Qualifier   |
| source           | 1..7522   | /organism="Homo sapiens"<br>/db-xref="taxon:9606"<br>/clone="pf01012"<br>/tissue-type="hippocampus"<br>/dev-stage="adult"<br>/note="vector:pBluescript II SK plus"  |
| gene             | 204..3626 | /gene="MEGF10"  |
| CDS              | 204..3626 | /gene="MEGF10"<br>/note="KIAA1780 protein gene encoding protein with multiple EGF-like-domains"<br>/codon-start=1<br>/product="MEGF10 protein (KIAA1780)"<br>/protein-id="BAB47409.1"<br>/db-xref="GI:14017777"<br>/translation="MVISLNSCLSFICLLCHWI<br>GTASPLNLEDPNVCSHWESYSVTV<br>QESYPHPFDQIYYTSCDILNWFKCTRHRVSYRT<br>AYRHGEKTMYYRKSQCCPGFYESG<br>EMCVPHCADKCVHGRCIAPNTCQCEPGWGGTNCS<br>SACDGDHWGPHCTSRCQCKNGALC<br>NPITGACHCAAGFRGWRCEDRCEQGTYGNDCHQR<br>CQCQNGATCDHVTGECRCPPGYTG<br>AFCEDLCPPGKHGPQCEQRCPCQNGGVCHHVTGE<br>CSCPSGWMGTVCQGPCPEGRFGKN<br>CSQECQCHNGGTCDAAATGQCHCSPGYTGERCQDE<br>CPVGTYGVLCAETCQCVNGGKCYH<br>VSGACLCEAGFAGERCEARLCPEGLYGIKCDKRC<br>PCHLENTHSCHPMSGECACKPGWS<br>GLYCNETCSPGFYGEACQIQSCQNGADCDSVTG<br>KCTCAPGFKGIDCSTPCPLGTYGI<br>NCSSRCGCKNDAVCS PVDGSC TCKAGWHGVDCSI<br>RCPSGTWGFGCNLTQCLNGGACN<br>TLDGTCTCAPGWRGEKCELPQDGTYGLNCAERC |

DCSHADGCHPTTGHCRCCLPGWSGV  
HCD5VCAEGRWGPNC5LPCYCKNGASCS PDDGIC  
ECAPGFRGTTCCQRICSPGFYGHRC  
SQTCPQCVHSSGPCHHITGLCDCLPGFTGALCNE  
VCP5GRFGKNCAGICTCTNNGTCN  
PIDRSCQCYPGWIGSDCSQPCPPAHWGPNCIHTC  
NCHNGAFCSAYDGECKCTPGWTGL  
YCTQRCPLGFYGKDCALICQCQNGADCDHISGQC  
TCRTGFMGRHCEQKCP5SGTYGYGC  
RQICDCLNNSTCDHITGTCYCSPGWKGARCDQAG  
VIIVGNLNSLSRTSTALPADSYQI  
GAIAGIILVLVVLFLALFIIYRHKQKGK5SSM  
PAVTTYTPAMRVVNADYTISGTLPH  
SNGGNANSHYFTNPSYHTLTQCATS PHVNNRDRM  
TVTKSKNNQLFVN LKNVNPGRGP  
VGDCTGTL PADWKHGGYLNELGAFGLDRSYMGKS  
LKDLGKNSEYNSSNCSLSSEN5PY  
ATIKDP5VLIPKSSECGYVEMKSPARRDSPYAEI  
NNST5ANRNVYEVEPTVS5VQGVF  
SNNGR5QDPYDLPKNSHIPCHYDLLPVRDSS5S  
PKQEDSGGSS5N5SS5S5E"

SEQUENCE (SEQ):

|      |             |             |              |             |             |             |
|------|-------------|-------------|--------------|-------------|-------------|-------------|
| 1    | aaagattgct  | ttcttctggt  | acgctgcttg   | aacgctaacc  | gcgttgattg  | gaacagattt  |
| 61   | tgtgtcttgg  | ctggcttgg   | gtgaagaccg   | gggagaaaag  | gttgcgctgc  | gattctcaag  |
| 121  | atctctggac  | ctggattatc  | attgcaaatic  | cattaaagaa  | ggagaagcaa  | gcggatttca  |
| 181  | gagaggttgt  | ttctcagaaa  | aaaatggtta   | tttctttgaa  | ctcatgcctg  | agctttattt  |
| 241  | gtttattgtt  | atgccactgg  | attgggacag   | catcacctct  | gaatcttgaa  | gaccctaattg |
| 301  | tgtgtagcca  | ctgggaaagc  | tactcagtga   | ctgtgcaaga  | gtcataccca  | catccctttg  |
| 361  | atcaaattta  | ctacacgagc  | tgcactgaca   | ttctaaactg  | gtttaaatgc  | acgcggcaca  |
| 421  | gagtcagcta  | tcggacagcc  | tatcgacatg   | gggagaagac  | tatgtatagg  | cgcaagtctc  |
| 481  | agtgttgctc  | tggattttat  | gaaagcgggg   | aaatgtgtgt  | ccccactgt   | gctgataaat  |
| 541  | gtgtccatgg  | tcgctgtatt  | gtcctcaaaica | cctgtcagtg  | tgagcctggc  | tggggaggga  |
| 601  | ccaactgctc  | cagtgcctgc  | gatggtgatc   | actggggctc  | ccactgcacc  | agccggtgcc  |
| 661  | agtgcaaaaa  | tggggctctg  | tgcaacccca   | tcaccggggc  | ttgccactgt  | gctgcgggct  |
| 721  | tccggggctg  | gcgctgcgag  | gaccgctgtg   | agcaggggcac | ctatggtaac  | gactgtcatc  |
| 781  | agagatgcca  | gtgccagaat  | ggagccacct   | gcgaccacgt  | cacgggggaa  | tgccgctgcc  |
| 841  | caccaggata  | caccggagcc  | ttctgtgagg   | atctttgtcc  | tcctggtaaa  | catggtccac  |
| 901  | agtgtgagca  | gagatgccct  | tgtcaaaatg   | gaggagtgtg  | tcatacagtc  | actggagaat  |
| 961  | gctcttgccc  | ttctggctgg  | atgggcacag   | tgtgtggtca  | gccttgcccc  | gagggtcgct  |
| 1021 | ttggaaagaa  | ctgttcccaa  | gaatgccagt   | gccataatgg  | agggacgtgt  | gatgtgcca   |
| 1081 | caggccaatg  | tcattgcagt  | ccaggatata   | caggggaaacg | gtgccaggat  | gagtgtcctg  |
| 1141 | ttgggacctg  | tggcgcttctc | tgtgtcgaga   | cctgccagtg  | tgtcaacgga  | gggaagtgtt  |
| 1201 | accacgtgag  | cggcgcatgc  | ctctgtgaag   | caggctttgc  | tggcgagcgc  | tgcgaagcac  |
| 1261 | gacctgtctc  | tgaggggctc  | tacggcatca   | aatgtgacaa  | acgggtgtccc | tgccacttgg  |
| 1321 | aaaacactca  | tagctgtcac  | cccattgtctg  | gagagtgtgc  | ctgcaagccg  | ggctggtcag  |
| 1381 | gactctactg  | taatgagaca  | tgttctctctg  | gattctacgg  | ggaagcttgc  | cagcagatct  |
| 1441 | gcagctgcca  | aaatggggca  | gactgtgaca   | gtgtgactgg  | aaagtgcacc  | tgtgccccag  |
| 1501 | gattcaaagg  | aattgactgc  | tctaccccat   | gccctctggg  | aacctatggg  | ataaactgtt  |
| 1561 | cctctcgctg  | tggctgtaaa  | aatgatgcag   | tctgtctctc  | tgtggacggg  | tcttgtactt  |
| 1621 | gcaaggcagg  | ctggcacggg  | gtggactgct   | ccatcagatg  | tcccagtggc  | acatggggct  |
| 1681 | ttggctgttaa | cttaacatgc  | cagtgcctca   | acgggggagc  | ctgcaacacc  | ctggacggga  |
| 1741 | cctgcacgtg  | tgcacctgga  | tggcgcgggg   | agaaatgcga  | acttccctgc  | caggatggca  |
| 1801 | cgtacgggct  | gaactgtgct  | gagcgctgcg   | actgcagcca  | cgcagatggc  | tgccacccta  |
| 1861 | ccacggggcca | ttgccgctgc  | ctccccggat   | ggtcagggtg  | ccactgtgac  | agcgtgtgtg  |
| 1921 | ctgaggggacg | ctggggcccc  | aactgctccc   | tgccctgcta  | ctgtaaaaat  | ggggcttcat  |
| 1981 | gctccccctga | tgatggcatc  | tgcgagtgtg   | caccaggctt  | ccgaggcacc  | acttgtcaga  |
| 2041 | ggatctgtctc | ccctgtgttt  | tatgggcatc   | gctgcagcca  | gacatgcccc  | cagtgcgttc  |
| 2101 | acagcagcgg  | gccctgccac  | cacatcacccg  | gcctgtgtga  | ctgcttgcct  | ggcttcacag  |
| 2161 | gcgccccctg  | caatgaagtg  | tgtcccagtg   | gcagatttgg  | gaaaaactgt  | gcaggaattt  |
| 2221 | gtacctgcac  | caacaacgga  | acctgtaacc   | ccattgacag  | atcttgtcag  | tgttaccctg  |
| 2281 | gttggattgg  | cagtgactgc  | tctcaacctat  | gtccacctgc  | ccactggggc  | ccaaactgca  |
| 2341 | tccacacgtg  | caactgccat  | aatggagctt   | tctgcagcgc  | ctacgatggg  | gaatgtaaat  |
| 2401 | gcaactcctgg | ctggacaggg  | ctctactgca   | ctcagagatg  | tcctctaggg  | ttttatggaa  |
| 2461 | aagattgtgc  | actgatatgc  | caatgtcaaa   | acggagctga  | ctgcgaccac  | atttctgggc  |
| 2521 | agtgtacttg  | ccgcaactgga | ttcatgggac   | ggcactgtga  | gcagaagtgc  | ccttcaggaa  |
| 2581 | catatggcta  | tggctgtcgc  | cagatatgtg   | attgtctgaa  | caactccacc  | tgcgaccaca  |
| 2641 | tcaactgggac | ctgttactgc  | agccccggat   | ggaaggggagc | gagatgtgat  | caagctggtg  |
| 2701 | ttatcatagt  | tggaaatctg  | aacagcttaa   | gccgaaccag  | tactgtctct  | cctgtgattt  |
| 2761 | cctaccagat  | cggggccatt  | caggagcatca  | tcattcttgt  | cctagtgtgt  | ctcttcctac  |
| 2821 | tggcattgtt  | cattattttat | agacacaagc   | agaagggaaa  | ggaatcaagc  | atgccagcag  |
| 2881 | ttacctacac  | ccctgctatg  | agggtcgtca   | atgcagatta  | taccatttca  | ggaacccttc  |
| 2941 | ctcacagcaa  | tgggtggaac  | gctaatagcc   | actacttcac  | caatcccagt  | taccacacgc  |
| 3001 | tcacccagtg  | tgccacatcc  | cctcacgtca   | acaacaggga  | caggatgact  | gtcacgaagt  |
| 3061 | caaaaaacaa  | tcaactgttt  | gtgaatctta   | aaaatgtgaa  | ccctgggaag  | agaggccctg  |
| 3121 | tgggggactg  | cactgggaca  | ttgccggctg   | actggaaaca  | tggcggctac  | ctcaacgagc  |



|      |             |             |             |             |             |             |
|------|-------------|-------------|-------------|-------------|-------------|-------------|
| 3181 | tccgtgcttt  | tggacttgac  | agaagctata  | tgggaaaatc  | cttaaaagac  | ctgggaaaga  |
| 3241 | attctgaata  | taattcaagt  | aactgctccc  | taagcagttc  | tgagaaccca  | tatgccacta  |
| 3301 | ttaaagaccc  | acctgtactt  | atcccgaaaa  | gctcagagt   | tggttatgtg  | gagatgaaat  |
| 3361 | cgccggcacg  | aagagattcc  | ccatatgcag  | agatcaataa  | ctcaacttca  | gccaacagga  |
| 3421 | atgtctatga  | agttgaacct  | acagtgagt   | ttgtccaagg  | agtattcagc  | aataatgggc  |
| 3481 | gtctctccca  | ggatccatag  | gacctcccaa  | agaacagtca  | catcccttgt  | cattatgacc  |
| 3541 | tgtcgccagt  | ccgagacagt  | tcactctccc  | ctaagcaaga  | ggacagtgga  | ggtagcagca  |
| 3601 | gcaacgacag  | cagcagcagt  | gaatgcaccc  | aaaggaccgc  | ttggtagcca  | ctggaaacct  |
| 3661 | ttccagaact  | gctgttttgt  | tcttctccat  | cctcaatttt  | gccactttca  | tgtgaatgtt  |
| 3721 | agtcaattcg  | gtgggcaatt  | tttggacatg  | aaccagaaa   | ctgaaagctg  | aggctgacac  |
| 3781 | ggactgtagg  | tgcTTTTTgt  | tcaggtggat  | tcgaaggagt  | tagagatgtg  | atttgccatt  |
| 3841 | gctgttagtt  | ttagaactat  | acccgtgaag  | catgacttat  | tgtaatgtgt  | tggctgaaa   |
| 3901 | catgaacttg  | cagaactccc  | tcggagacgc  | agggttcagt  | ggacattggg  | attgttgctt  |
| 3961 | gaaaaattaa  | aatttgaata  | ttttctctct  | catttgcata  | atacagctct  | acctaggatt  |
| 4021 | gtacagttta  | ccataaaaatt | tacttcatga  | aagtgggaat  | cactgaacat  | gtagaagaca  |
| 4081 | aggaacatat  | tgtaaactcc  | tgattcttaa  | ctttattcaa  | ctggactcag  | aattgtaggg  |
| 4141 | ataatatgaa  | tgcaggagga  | aacattctgt  | caggcggtat  | gactggacag  | actttgaata  |
| 4201 | tactctaaaa  | gtggacagaa  | aattttacgaa | aatcttagat  | tttgtttaga  | atgagaaaat  |
| 4261 | atacaattag  | aattatttta  | gaaatagtga  | gaagtattgc  | agaagtcaat  | acacaaatgt  |
| 4321 | gccaggcaga  | ggtggttttc  | tctgtttgac  | tctcaaccac  | cttcagatct  | atgacattat  |
| 4381 | tctgatcact  | ggctccatca  | tacataattca | ccacttgaga  | ttcataacat  | atcaatagtt  |
| 4441 | atttcataaa  | tatagaaaatg | aaataatttt  | atttttgaca  | gactggatgg  | aatgagtgtg  |
| 4501 | taatgattga  | taaagggtgt  | aaattttaaa  | tgcaagatga  | cgcttacgtt  | ctgtaaacca  |
| 4561 | ttagtaatac  | atgctgtaat  | atagaattag  | tggaaacattt | tgattaatct  | ttccctagaa  |
| 4621 | gtgactgaaa  | tattttttgt  | catatttgag  | aaagggaact  | ttccttttat  | taattgtcaa  |
| 4681 | tttagagaaa  | ctatgcttaa  | gctggtcttt  | tgcatgtcta  | atgtgacatg  | taccaacttt  |
| 4741 | ttcatttaatt | tgtattttcca | tttttaaatt  | gcatatttcta | tgttttgtag  | tgtttgattt  |
| 4801 | gttaatgaaa  | aaatattata  | tgttcggtat  | tccttgattt  | attgccactt  | atcttttgct  |
| 4861 | tgataaaaaat | gcgttgttct  | tttttctttt  | ggagggacaa  | gatgaaaata  | tataatttga  |
| 4921 | attgattaaa  | attggctcgtt | actaaaatag  | tatagtaacc  | acaagtgatt  | ggcttataaa  |
| 4981 | tgaagttagaa | atgcttttta  | atattccaaa  | atagagttcc  | ttttgatctg  | ttggtgctga  |
| 5041 | gccttggtta  | aaccagggag  | aaggggagca  | gaaaggaaac  | ggtgttactg  | atgagtacca  |
| 5101 | cagactctac  | ttaaaaaaaa  | ctctcattat  | ggtgatcata  | gaattgacca  | tccaaactgg  |
| 5161 | gacactcttg  | agagtaaatg  | gagggcatta  | ttataaatta  | tcttgtaatg  | aacttaaatac |
| 5221 | tggactgttc  | caggcaaacc  | agacttatct  | tgcaatatga  | gaatgctgac  | acaatgcagg  |
| 5281 | aaagccagtt  | tcccttttgt  | tgatctactt  | gaccaagcaa  | aggggctgaa  | aaactgaata  |
| 5341 | aggaaacaac  | tttataagag  | aaacagtggt  | cttcaatctt  | ttaaagacat  | gaaatcctat  |
| 5401 | atggcattct  | gtctcagtga  | gtcagttaac  | aaatacgtat  | gtgcaaccct  | tctgctagta  |
| 5461 | gtgcacataa  | gtgattatcc  | ctggcaggga  | tcgagttgga  | atatccagtt  | atttcatgtc  |
| 5521 | acacatcggc  | acgtatgatg  | gtgggtttgt  | cagatggata  | atacagcaga  | cacacttaga  |
| 5581 | acactcactg  | cactgggtggc | tgttcatttt  | gaggaactcc  | aaagtcaatt  | caaggaaata  |
| 5641 | aggaatgacc  | tggaaacaggc | cttagaaaaca | attgattttat | tccaatagtt  | aaccactggc  |
| 5701 | tggctcccaa  | ctctagggtga | taggcatcta  | attgagacat  | gtgtgagtca  | atagccatcg  |
| 5761 | gggtccctttt | tggtgagaag  | aaataagaca  | ttttctccct  | tttaaagatc  | atctcttcag  |
| 5821 | agctgcattg  | ttaaatttgaa | tttccccatt  | gttctctaga  | gatagggtcaa | aaaaacaact  |
| 5881 | gtttcttttt  | ttattccact  | attaataaat  | tgaaagcttg  | ctgtacaaag  | gcaacagcaa  |
| 5941 | cttaaaaaag  | aaaatctgga  | caatagatgt  | ggactccaag  | gggccactgc  | catcccttcc  |
| 6001 | tgtgtgctct  | ttttgacaag  | taaatcactc  | cagtttagctg | atgtcctgaa  | atgacatttg  |
| 6061 | tacctgaggc  | aattcttatt  | agagcttact  | caggactttt  | caaacatcag  | gacacacata  |
| 6121 | gcaattttgga | tgttgtaaaa  | ctactgttta  | cattttggaga | ggtcttggtg  | agaaggcaaa  |
| 6181 | atttttctaaa | tattttgata  | tcaatttggg  | taaagaaagg  | aatacttttg  | ttagaatgag  |
| 6241 | aattaaagag  | aaaacaaaag  | aatcttgcca  | aaattttccc  | tctgaaatta  | caaactttag  |
| 6301 | gaacttttcc  | agatgtatcg  | aattttaattt | gacattgatg  | tcactcttca  | taccaagtga  |
| 6361 | atctattctc  | atgaactttg  | aacccatggt  | aattttggac  | cctataacta  | atttcctttc  |
| 6421 | tgtatcccat  | ggtggtcaaa  | gcagtaaagg  | agtttagaga  | tacttagcaa  | atgccattca  |
| 6481 | tatagtattt  | ttagccaaac  | tattattcct  | tccttagaag  | ttgatgtatc  | agaaaattta  |
| 6541 | taagtatttg  | gtattttatg  | ataagtacat  | cagaacttgc  | ccacatgata  | actcatgttt  |
| 6601 | gctttaataa  | ggaaaatatg  | ttatttgagg  | ttaatatat   | gatagtgagg  | aacattttac  |
| 6661 | aattattttac | agtgtttgac  | attaattgct  | gtcctatgat  | acagtagtaa  | tgtaaacaga  |
| 6721 | gagtttaata  | tgtttagcat  | tatctttatg  | gaatttatat  | tcaccaattt  | caggaaaacc  |
| 6781 | aaccatagaa  | ctttataata  | ggacagttat  | aacctgactg  | aaccaaattc  | atcagcataa  |
| 6841 | tgactgtgtg  | ttctgagaaa  | tgcaaaacaac | tttatgaaag  | tgtggtcatc  | tgaccccca   |
| 6901 | aagagccaca  | gtagctgtca  | attattgaga  | tttttaaattg | gtaaatattg  | caaaagttaa  |
| 6961 | acaagggtac  | cacaatatca  | tttataattg  | atgtcaatat  | tagtagtcca  | ctttatgctt  |
| 7021 | aaaataaaat  | cataaaaaat  | caagtttttaa | ggttatatatt | tttaaaataa  | tgaattttct  |
| 7081 | atctctaggc  | aacatgtctt  | tattattcaa  | gctgaatgtt  | taagagagat  | tttggcttta  |
| 7141 | aaggcttcat  | catgaaagt   | tacatgcata  | tgcaagtgtg  | aattacgtgg  | tatggatggg  |
| 7201 | tgcttgttta  | ttactaaag   | atgtacagca  | aactgcccgt  | ttagagtcct  | gttaatatgt  |
| 7261 | atgtcctaac  | actgggtctg  | cttatgctag  | ttcagtattt  | ggcttaggct  | tcattttgat  |
| 7321 | atgactgaat  | tgcaatctat  | atttttataa  | agaaaaatca  | aaatacatta  | gggtttactt  |
| 7381 | ggttgtggca  | aacaaacaaa  | aaatgctttc  | agtgttgaaa  | tatctctatt  | ttccaaagat  |
| 7441 | gataaacttt  | catcgttctt  | agcctacatt  | gtcatttata  | tcaccaaatg  | agtttataga  |
| 7501 | ttaatgcaat  | aaactttcta  | at          |             |             |             |